DOCUMENT RESUME

ED 115 131 FL 007 335

AUTHOR Iannucci, David; Dodd, David
TITLE Surface Syntactic Structures and Linguistic Memory: Some Experimental Evidence. San Jose State Occasional Papers in Linguistics, Vol. 1.
INSTITUTION San Jose State Univ., Calif.
PUB DATE 3 May 75
NOTE 10p.; Paper presented at meeting of the California Linguistic Association (5th, May 3, 1975)
AVAILABLE FROM Linguistics Program, San Jose State University, San Jose, California 95192 ($4.00 per volume)

EDRS PRICE MF-$0.76 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Deep Structure; Grammar; Language Research; Linguistics; *Memory; Nominals; *Psycholinguistics; Recall (Psychological); Retention; Semantics; Sentence Structure; *Surface Structure; *Syntax; *Transformations (Language); Verbs
IDENTIFIERS *Critical Pairs

ABSTRACT

This paper describes and gives the results of a psycholinguistic experiment investigating the impact of certain surface syntactic structures on the perception and memory of language. The basic assumption is that the content of an utterance must be its most salient aspect in memory. The for of an utterance, its surface grammar and phonology, must have some salience for memory, especially short-term, but nothing like that of content. Amy meaningful grammatical distinctions of markers should, therefore, be more salient in memory than meaningless ones. This experiment focuses exclusively on different varieties of one surface distinction: sentences whose underlying predicate surfaces as a verb versus those where the same predicate surfaces as some nominalization of that verb. The experiment consisted of administering to the subjects pairs of synonymous sentences, differing in the above manner and referred to as Critical Pairs (CP's), and five other types of sentences, included for comparing recognition performance on CP's to performance on other kinds of linguistic memory. The conclusion is that the surface syntactic differences manipulated here have little, if any, salience for memory, and that it is unlikely for such surface differences to have any semantic consequence. Such conclusions, however, do not necessarily generalize to other surface features. (Author/TL)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICRO FICHE ONLY HAS BEEN GRANTED BY
DAVID IANNUCCI

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED115131

FL007335

Surface syntactic structures and linguistic memory: some experimental evidence

David Iannucci and David Dodd (University of Utah)

INTRODUCTION. We are interested in the impact of certain kinds of surface syntactic structures on the perception and memory of language. Regarding linguistic memory, we take as a basic assumption that the content (meaning) of an utterance must be its most salient aspect in memory (or else the basic function of language would be quite strange). The form of an utterance--its surface grammar and phonology--must have some salience for memory (esp. short-term memory), but nothing like the salience of content. It follows that any meaningful grammatical distinctions or markers should be more salient in memory than meaningless ones.

Part of the direction we've taken in exploring this matter has been inspired by two very different points of view. First, some very interesting psycholinguistic experiments have been done by Sachs (1967) and by Bransford and Franks (1971) which show that people recall the interpretation of a sentence far better than its syntactic form--even to the extent that certain kinds of entailed information are just as salient in memory as explicitly encoded information. Subjects in the Bransford and Franks study even reported having heard whole sentences which had never been presented to them in the experimental materials but were clearly entailed by sets of sentences they had actually heard. We are not particularly concerned with the problem of entailed versus explicitly encoded information, however. Instead, we wanted to construct specific grammatical hypotheses about particular surface structures and memory.

This brings us to the second source of our 'inspiration': Benjamin Lee Whorf. Whorf made some very strong claims regarding the impact of different surface structures across languages on the understanding of content by speakers of those languages. One of his favorite examples (1956. 147, 215, and elsewhere) concerns the assertion that (esp.) spatial and temporal events in nature were handled quite differently in Hopi and English. Specifically, he contended that Hopi would typically refer to these events ('eventings') with verbs and that English would use nouns, and further that this difference caused (let alone merely reflected) a very different understanding of those events by speakers of the two languages. So, presumably, the Hopi and I look out over the ocean at the same natural phenomenon. He says something in Hopi like "The ocean waves (is waving)", and I say "There are waves on the ocean". Thus, à la Whorf, our 'views' of the event are quite different: the Hopi understands it as the true continuous motion that it is, and I chunk up ('objectify') the peaks ('peaking'?) of that motion as if they were discrete things. Whorf clearly meant such interpretations quite strongly. Whereas most linguists would now see such a difference as merely surfacing the same underlying predication (WAVE) in one language as a surface noun and in the other as a surface verb, Whorf's claims must entail, in modern terms, either some difference in the underlying structure, or the possibility that the transformations responsible for the different forms in fact alter meaning. The latter is plausible, especially if we soften the strength of Whorf's assertions considerably. The surface difference between nominal versus

verb representation of the same underlying conceptual category could have slight potential, possibly pragmatic, semantic consequences.

In any case, you don't need two languages like Hopi and English to examine these matters. English alone provides ample opportunities for encoding an underlying predication as a surface verb or the corresponding noun, i.e. that verb's nominalization. If such differences are crucial to meaning across languages, they ought to be so within one language too. Thus we have the working grammatical variable for our experiment: what are the consequences for memory of this one type of syntactic difference in English--where an underlying predicate surfaces as a verb versus its nominalization (along with other necessary syntactic adjustments that follow)? If such syntactic differences are meaningful, given our initial assumption, then we expect them to have some impact on the remembrance of the event, hopefully comparable to the impact of other more obvious meaningful differences.

EXPERIMENTAL DESIGN. The design of the experiment was as follows. We took two passages on relatively interesting topics and also written in relatively breezy and entertaining styles. We edited them in ways to be described below; and then recorded them on tape. Each passage was about 7.5 minutes long. 80 subjects sat and listened to the two recorded passages--a total of 15 uninterrupted minutes--with the instruction to simply listen carefully (nothing else). We then gave them a two-part sentence-recognition memory task based on the two passages--the first part on the first passage, and the second part on the second passage. Thus a rather busy time-gap was introduced inbetween their listening to a passage and our request for them to recall certain things about that passage: for Passage A, the time-gap was filled by listening to Passage B; for Passage B, the time gap was filled by the memory test on Passage A.

The memory tests each consisted of a list of sentences--51 for Passage A and 47 for Passage B. The subjects were instructed to pick out the sentences on the list(s) which they actually heard in the passage(s), but only those that they thought they heard exactly, word-for-word. They were to reject any sentence they hadn't heard or that had been altered in any way. The sentences they were given were of six types.

The first type consisted of sentences involving our target variable. We will refer to these as 'Critical Pairs' (CP), and some examples are as follows:

V	1)	He would risk his life to defend her.
Nom.	2)	He would risk his life in her defense.
V	1)	They did so to escape tedium.
Nom.	2)	They did so as an escape from tedium.
V	1)	We assume, moreover, that...
Nom.	2)	Our assumption, moreover, is that...
V	1)	He really needs his wife.
Nom.	2)	He has a real need for his wife.
V	1)	This study supports the common view that...
Nom.	2)	This study lends support to the common view that...

V 1) We hoped to learn if they used body language to communicate cues of...

Nom. 2) We hoped to learn if their use of body language communicated cues of...

In the original passages, there were a number of points (26 in Passage A and 22 in B) where one member of a CP occurred. At each of these points, either the verb-type (V-type) or the nominalized type (N-type) occurred. We made up two versions of each passage (a & b), so that for each CP, when the N-type occurred in Version a, the V-type occurred in Version b, and vice-versa. The choice of V versus N-type at any one point was distributed randomly across the 26/22 CP's in the passages. Half of the subjects heard Version a of the two passages, and half heard Version b. We similarly counterbalanced the test forms to control for V-type versus N-type and for whether the sentence was identical to the one heard or changed (V to N-type or vice-versa). We made up two different test forms for each passage so that--on one test form--half of the N-types and V-types were the same on the test as on the original passage and half were changed. On the other test form, the items were the opposite. An example should make the upshot of the above clear. Consider four types of subject for Passage A. Two of them hear (on Version a) the (V-type) sentence "He would risk his life to defend her". One of the two has, on his test form, the same sentence, and the other is presented with (N-type) "He would risk his life in her defense". The other two subjects hear (on Version b) "He would risk his life in her defense"; one has that sentence on his test form, and the other has "He would risk his life to defend her". What this further means for any one subject, whatever combination he was presented with, is that his test form included (e.g. for Passage B) 11 sentences drawn from CP's that he really did hear word-for-word. For the same subject, there were also on the test form 11 sentences drawn from CP's that he didn't hear--instead, in these cases, he really heard the other member (type) of the CP. Without such counterbalancing, one cannot consider each critical set of sentences with confidence, since (a) the content contained in a sentence might be relatively salient or memorable, thus producing a greater than average percentage of 'yes' responses, or (b) the verb form might be more stylistically natural (or whatever) so that it generated more 'yes' responses. Since the complete counterbalance means that every sentence will (across 4 groups of subjects) be presented as an N-type or V-type and (for each form of presentation) tested as N-type or V-type, each sentence's results can be examined appropriately.

The remainder of the sentences on the test form, for the recognition task, were of five types and were the same for all 80 subjects on all test forms and on both versions of the two passages. Each of the 5 types either introduced a sentence the subjects had not heard at all, or a sentence that was an altered version of one he actually heard. All were designed as controls for comparing recognition performance on CP's to similar performance on other kinds of linguistic memory.

Type I: In our editing of the original passages, we removed whole sentences and paragraphs. We, of course, had to do so in ways we thought were not stylistically or otherwise disruptive to the passage. We then

presented some of these removed sentences on the test forms. Type I were a group we felt to be strongly entailed by the passage as a whole, i.e one didn't need to hear the sentence to get the information, if he heard the whole passage. What we anticipated here was an effect similar to the Bransford and Franks (1971) study: a number of subjects ought to think they really heard these sentences, even though they were merely entailed by the passage.

Type II: These sentences (removed from original passages) were also never heard by any subject. They were not entailed but they were at least highly consistent with the passage as a whole. Not only was it very plausible that such sentences could have occurred in the passage, the original authors had included these sentences (as with Type I). But since they added essentially new, even though consistent, content we expected that subjects would probably not be fooled and do a pretty fair job of rejecting these as sentences they didn't hear.

Type III: These sentences, as well as those of Types IV and V, rather than introducing new sentences, introduced alterations on actually heard sentences. For Type III, we made minor meaningful lexical changes on original sentences, e.g.:

HEARD: 1) He adores his wife.
ALTERED: 2) He cherishes his wife.

These were all quite slight and non-disruptive to the context or flow of the passages, i.e. they were designed such that if the altered sentence had occurred in the original passage, the change would have been of minimal consequence--but even so, the change was a meaningful one. We anticipated here that, given the level of difficulty of the memory task, these minor changes would be unnoticed to a fairly large degree.

Type IV: For this type, on the other hand, we altered original sentences considerably and, again, the changes were semantic ones. The changes, however, were far from outrageous. They all still could have fit quite well in their passages; e.g.:

HEARD: 1) I don't like you anymore.
ALTERED: 2) I never did like you.

In the above example, the passage had a woman bitterly condemning her thoughtless husband for the way he treated her. She rattles off a long list of grievances and then winds up on this emotional line. Sentence (1) versus (2) makes a big difference in meaning, but either one would be highly appropriate in the given context, especially since the matter at hand was not a topic of either prior or further discussion--she just 'up and said it'. Our prediction here was that subjects would do a good job of recognizing these semantic changes.

Type V: In this type, unlike Types III and IV, we made alterations on actually heard sentences that we deemed non-meaningful--or as close as we could come to it; e.g.:

HEARD: 1) They touched themselves more frequently than they touched each other.

ALTERED: 2) They touched themselves more frequently than they touched one another.

For this type, we predicted the highest level of sentence recognition error; we strongly doubted subjects' ability to recognize that these sentences had been changed, since the changes were quite minimal in form--usually one or two words--and non-semantic.

In sum, then, our subjects went through a list of sentences (for each passage) which involved the six different types of manipulation (randomly mixed) described above, and made their decisions on what they thought they heard. In addition to subjects classifying the sentences as heard ('yes') or not heard ('no'), we asked them to also indicate their confidence in each decision, on a scale of 1 (guessing) to 5 (absolutely certain). Since these confidence values are consistent with but add nothing of interest to the overall interpretation of results tabulated without them, they are not reported below.

RESULTS. The results are as follows. Subjects' performance on sentence recognition of CP's alone is quite revealing. There were essentially two conditions for CP recognition, each in effect for half of the CP's for any one subject. In one condition the CP sentence on the test form was identical to what was heard in the passage (SAME condition, correct answer='yes'). In the other condition, the test form presented the other sentence in the CP pair from what was actually heard (CHANGED condition, correct answer='no'). In Passage A, subjects correctly identified sentences they had heard, in the SAME condition, 69.7% of the time. In condition CHANGED, however, they continued to assert recognition of the CP sentences they hadn't heard, and % correct scores dropped to a mean of 30.7%. Note that 30.7% recognition accuracy is equivalent to 69.3% recognition error, thus the rate of false identification in condition CHANGED matches the rate of correct identification in condition SAME almost exactly. Subjects equally reported having heard the CP type on the test form whether they had or not and the difference between this 'yes' response in the two conditions (69.7% versus 69.3%) is not at all significant.

In Passage B, the percentages are different, but the results are the same (we discuss overall differences between Passages A and B below). In the SAME condition, subjects correctly identified heard sentences at a 76.3% level of accuracy. In condition CHANGED, recognition accuracy drops to 26.2%, equivalent to a recognition error mean of 73.8%. Again, the difference (76.3% versus 73.8%) between 'yes' responses, whether correct or not, is not significant. Given the CP results alone, then, it could hardly be argued that there was any salience in memory at all for the difference between N-types and corresponding V-types manipulated in this task. The almost exact match of % recognition error with % recognition accuracy in the two conditions is particularly striking.

The interpretation of results from Types I-V sentences is considerably more complicated. Those results are summarized below. Statistical significance between adjacent percentages are parenthetically indicated by 'sig.', non-significant ones by 'n.s.'. All reports of significance here

entail $p < .05$ or better. It is important to note that, unlike the CP manipulation, Types I-V always entailed a changed sentence on the test form, thus the correct answers for these were always 'no' (exact sentence not heard). For this reason, we include (for comparison) CP results only from the CHANGED condition. Percentages are listed in ascending order of recognition accuracy.

PASSAGE A

% correct response

(CHANGED)

Type:	V	CP	III	I	IV	II
	28.4 (n.s.)	30.7 (n.s.)	35.6 (sig.)	49.6 (sig.)	68.4 (sig.)	80.6

(V vs III is sig.)

PASSAGE B

% correct response

(CHANGED)

Type:	V	CP	III	I	IV	II
	17.0 (sig.)	26.2 (n.s.)	27.8 (sig.)	49.0 (sig.)	81.8 (n.s.)	84.4

Note that there are some clear differences between categories across the two passages. In fact the differences between Passages A and B, for CP's, Types III, IV and V, are all significant. We cannot give a sure explanation for these differences, however, the two passages do differ in two (possibly explanatory) ways. First, the busy 'gap' between hearing the passage and taking its test was different (see above), thus affecting the nature and perhaps difficulty of the memory task (we have no way of measuring such differences). Perhaps more importantly, the assembling of the two passages, along with the various manipulations of sentence changes for the test forms, were done by one investigator (D.D.) for Passage A, and by the other investigator (D.I.) for Passage B. How to accomplish the various manipulations was by no means a straightforward simple-minded task and the strategies we followed differed in certain respects, e.g. D.I. had a tendency to manipulate surface form somewhat less than D.D. over Type III, IV and V, but not for the CP's. In any case, however the cross-passage differences are explained, they are not inherently interesting, especially when compared to the fact that, despite significant cross-passage differences, the results within passages are distributed essentially the same and bear out initial predictions regarding Types I-V quite well.

The distribution of results across the six sentence categories had several interpretive functions. First it is quite clear that the memory task was such that the variety of manipulations varied in salience for memory from quite easily remembered to barely memorable at all.

Second, in this kind of experimental setting, there is always a danger of subjects falling into a 'yes' answer strategy, and of course our CP results (taken alone) weigh heavily on a high rate of 'yes' responses. Types I-V, however, always entailed a correct answer of 'no' and, where predicted (Types IV and II), subjects shifted to 'no' answers at a higher rate in general than that of the 'yes' responses for the CP's in either condition. Given this data, it is highly doubtful that such a 'yes' answering strategy affected the results.

Finally, we can attempt to compare CP results with Types I-V results. Note, however, that this must be done with great caution, since the 5 control types are not strictly comparable in design to the CP's. Most importantly, Types I-V were not counterbalanced with correct 'yes' answers on the test forms (i.e. no SAME condition), and further each one of Types I-V consisted of only 5 items for each passage (cf. 26 CP's for Passage A and 22 CP's for B). We can however make some tentative observations, regarding CP's versus Types I-V, with appropriate cautions expressed. Before doing so, however, we can briefly sum up the Type I-V results (see chart above).

The Type V synonymous lexical changes, as predicted, slipped by the subjects attempts at accurate recognition more than any other type (28.4% and 17.0% recognition accuracy for A and B respectively). Type III minor semantic lexical changes also slipped by more unnoticed than noticed, but subjects did significantly better on these (35.6% and 27.8%) than for Type V. Type IV, the only other sentence-alteration type (vs. whole new sentences of I and II), which consisted of more major semantic lexical changes, caused accuracy of recognition to jump well up into the accurate recognition range (68.4% and 81.8%)--again, no surprise.

Of the two wholly new sentence types, Type I, the highly implied-by-passage variety, introduces 50-50 chance results for memory (49.6% and 49%). This was clearly a confusing category and given studies such as Bransford and Franks (1971), hardly a surprise. It should be mentioned, however, that the Type I results for Passage B are really more of a mess than a clear indication of random memory: the range of correct responses varies from 16% to 71%, so even though the mean is 49%, it is probably not accurate to report this as a true random category for Passage B. Type II, the other variety of newly-introduced sentences is clearly the most recognizable of all: subjects uniformly rejected these at a very high rate (80.6% and 84.4%).

The question now is where do the CP results really fit with regard to the other 5 types? If we consider their place on the above chart, in Passage A, recognition of CP's seems to fall inbetween Types V and III (significantly different from neither, despite the fact that the V versus III differences are significant). In Passage B, CP recognition must be pegged with that of Type III (n.s. from Type III, but significant when compared to V). We might conclude from this (given the overall caution above) that the CP differences have a salience for memory roughly in-between that of III and V, and possibly closer to the slight semantic changes of III, thus lending potential support to a view that the N versus V surface differences do in fact entail very slight semantic differences. If taken seriously, this is somewhat surprising, given the total lack of salience indicated by the CP results taken alone, which would seem to indicate a mere formal difference more comparable to the Type V differences. There are a variety of possible explanations for this apparent discrepancy--all of which are highly speculative. The first is, of course, that for reasons expressed above, CP results are not strictly comparable to Types I-V results in general.

Another possible explanation entails some confusing results on particular CP items. In Passage B, for example, 17 of the 22 CP's showed a strong tendency to be correctly identified when heard and also to be equally incorrectly accepted when the opposite type of the pair was heard:

80.7% accuracy in the SAME condition; 18.7% in the CHANGED condition (=81.3% error). This of course pegs CP performance for the 17 items almost exactly on a par with recognition performance for Type V (17.0%) sentences. The other 5 CP's fell considerably out of this pattern. For these 5, in the CHANGED condition, subjects averaged just slightly better than chance (vs. 18.7% for the others) in correctly rejecting the test sentences. Something was more obvious about these 5 and acted as an aid to recognition. More surprisingly, in the SAME condition for these 5 sentences, accuracy dropped considerably (also) to barely better than chance recognition (vs. 80.7% for the others). (Similar adjustments for problematic items in Passage A yield essentially similar results to the above). The latter is particularly confusing to us, but the increased salience in the opposite condition may be a contextual or stylistic problem. We have some ideas, in this regard, about some of the 'offending' sentences, but they are far too speculative to merit further discussion. Of course, removing data or test-items after-the-fact is, in most cases, a bad experimental procedure. We speculate on the above possibilities only because of the rather confusing data which could indicate something faulty in these 5 items, which we simply cannot put our finger on at this time.

A final attempt to cope with the discrepancy at hand resulted from our observation that the purely surface (physical) adjustments on Type V sentences were quite minimal--usually involving only one or two words. The CP differences, however, tended to be more radical in most cases. For example:

HEARD: 1) He would risk his life to defend her.
ALTERED: 2) He would risk his life in her defense.

To get from (1) to (2), a phonological adjustment is made: 'defend' becomes 'defense'; 'her' moves from after to before the stem (thus taking on a different surface function); the clause in (1) is marked by 'to' and the clause in (2) by 'in'. And this is all pretty much a case of minimal CP change between the two (compared to other CP's). Despite the fact that such purely physical surface adjustments should have little salience in memory, they must have some, especially in relatively short-term memory, and this variable could work, at least to a small degree, as an aid to memory--more so for the CP's than the Type V sentences. If the two types were equal in this regard, we might expect the level of error on the CP's to be closer to Type V sentences than actually occurred in this study. In order to tease out any traces of such a pure physical difference effect, we attempted to devise a numerical index of surface difference between pairs of sentences and then test for correlations between this index and overall accuracy within the CP groups (the largest groups by far). We were able to find no such correlation. The problem, however, could be in the measure itself, which was essentially based on counting morphemes not shared by the two compared sentences (including process morphemes), with an additional (reduced) weight assigned to allomorphic differences. There is of course no precedent at all for such a measure (let alone a theoretically principled one), and a great deal of arbitrariness permeated the procedure. (It only takes a few minutes of attempting such a measure to get a strong feel for this.) All we can say, then, is that if the effect

is there, we cannot measure it and we cannot be sure why not.

In sum, then, the results of this experiment point to an overall conclusion that the surface syntactic differences manipulated here have very little, if any, salience for memory. If our initial assumption regarding the relative salience for memory of content versus pure form is true--and there is no evidence whatever to the contrary--then we must further conclude that it is highly unlikely for such surface differences to have any semantic consequence whatever. Note, however, that results such as these do not necessarily generalize to other surface features. Whereas several years ago, most generative linguists might have (rather cavalierly) rejected the possibility of surface differences (features) having any semantic consequence at all, current developments seem to be more open to such considerations; however, it is also clear that these considerations must be approached on a (surface) feature-by-feature basis.

Finally, we would like to suggest that the experimental method used here--particularly the exploitation of manipulations of linguistic memory--is a particularly fruitful one which can be refined and adapted to suit a number of interesting questions, especially those where sensitivity to slight semantic differences is crucial, such as various problems within the general question of synonymity and certain areas of pragmatics. Peter Ladefoged (1975) issues a plea for syntacticians to quantify and experimentalize their discipline, as has been done in other areas of Linguistics, rather than to continue to make crucial but untested assumptions on matters, such as syntactic synonymity, which can often be nuclear to an analytic argument. We acknowledge the fact that syntactic (let alone semantic) questions are not as amenable to such methods as are, for example, phonetics and phonology. However, the current state of the syntactic art shows virtually total disregard for such concerns, and remedies seem to be in order, especially when one considers that current syntactic/semantic research is in a position to generate an enormous number of interesting experimental hypotheses.

REFERENCES

Bransford, J.D., and J.J. Franks. 1971. The abstraction of linguistic ideas. Cognitive Psychology 2.331-350.

Ladefoged, Peter. 1975. Phonetics as the forerunner of linguistic science. Parasession on Research Methodology, Fifth California Linguistics Association Conference (San Jose).

Sachs, J.S. 1967. Recognition memory for syntactic and semantic aspects of connected discourse. Perception and Psychophysics 2.437-442.

Whorf, B.L. 1956. Language, thought and reality: selected writings of Benjamin Lee Whorf, ed. by John B. Carroll. Cambridge, Mass.: M.I.T. Press.